Evaluating anti-corruption agencies: learning from the Caribbean

Article (Accepted Version)

Evaluating Anti-Corruption Agencies: Learning from the Caribbean

Elizabeth David-Barrett[1], Aoife Murray[2], Johanna Polvi[3] and Richard Burge[4][5]

Abstract

Efforts to evaluate the impact of anti-corruption programmes face numerous difficulties related to the complexity and hidden nature of corruption, the political sensitivity of the topic, and the ability of corrupt networks to respond flexibly to interventions. This makes it difficult to measure changes in corruption levels and problematic to attribute them to interventions. This article shows how a theory-based evaluation that builds on academic research can elaborate a set of intermediate outcomes for the evaluation of anti-corruption programmes, and showcases several purpose-built tools for evaluating different aspects of capacity building. It also draws on learning from two evaluations of Anti-Corruption programmes in the Caribbean to demonstrate how anti-corruption theory is being translated into law enforcement practice in two ways: (i) through economic models of criminals as rational actors whose behaviour can be changed through incentives and disincentives; (ii) and through social norms models which argue that reducing corruption requires deeper social change that resonates with or adapts local norms.

Keywords: anti-corruption agencies, evaluation, Caribbean, capacity-building, social norms

[1] University of Sussex

[2] ITAD

[3] Independent Consultant

[4] ITAD

[5] Acknowledgements: The evaluations reported here were conducted by ITAD and commissioned by DFID. The authors would like to thank the independent reviewers and also participants in the Development Studies Association conference 2019 for their very helpful feedback on the paper

Introduction

Efforts to evaluate the impact of anti-corruption programmes face numerous difficulties. First, corruption is a complex social phenomenon which is difficult to detect or measure. Second, anti-corruption work is highly politically sensitive since it threatens to challenge those in power and disrupt political settlements, meaning that concrete achievements may not be observable within a single programming period and creating a need to ensure that intermediate outcomes – including capacity-building - are properly valued. Third, corrupt networks often respond flexibly even to well-designed interventions, again inhibiting our ability to trace the relationship between programmes and outcomes. Overall, it is difficult to measure changes in corruption levels and even more problematic to attribute changes to particular interventions.

This article follows Ravallion (2009) in arguing that a theory-based evaluation can work in such circumstances. A theory-based evaluation can build on academic research about what works in tackling corruption to elaborate a set of intermediate outcomes that can be evaluated. By teasing out underlying assumptions in a theory of change, it is possible to provide reasoned and evidence-supported judgements on performance, as well as to design recommendations for medium-term reform. Drawing on theory and research about corruption control, we particularly emphasise the importance of different kinds of capacity in Anti-Corruption Agencies (ACAs). We introduce three purpose-built tools for evaluating capacity building in ACAs: organisational capacity assessment, network analysis and policy tracking.

The article illustrates the value of these tools by drawing on learning from two theory-based evaluations of the five-year (2015-20) UK DFID Caribbean Anti-Corruption Programme (CACP). The CACP was designed to reduce crime through support for law enforcement agencies and financial investigations in Jamaica and the Eastern Caribbean. We undertook two mid-programme evaluations: one evaluated the Jamaica-based components, namely support for the Major Organised Crime and Anti-Corruption Agency (MOCA), the Financial Investigations Division (FID), and the Independent Commission of Investigations (INDECOM); the other focused on DFID support to the Asset Recovery Unit (ARU), a unit within the Regional Security System (RSS) in Barbados which supports financial crime investigation and asset recovery in Barbados and six islands of the Organisation of Eastern Caribbean States (OECS). The sought impact of the CACP is better governance in Jamaica and the Eastern Caribbean due to a reduction in corruption and organised crime, while the intended outcome is that public agencies in Jamaica and the Eastern Caribbean jurisdictions combat corruption and serious organised crime more effectively.

This article proceeds as follows. First, two main theories on how to reduce corruption are introduced and critiqued. Second, some of the difficulties of measuring corruption and evaluating anti-corruption interventions are discussed. Third, the concept of theory-based evaluation is explained and related to the specific problem of evaluating anti-corruption agencies. Fourth, three purpose-built tools are introduced and explained. Fifth, our conclusions share learning to inform evaluations of similar agencies.

Anti-corruption theory

Anti-corruption interventions are typically based on one of two models of human behaviour. One model, rooted in economics, characterises humans as rational actors who make calculations about how to act, weighing up the costs and benefits of different courses of action (Rose-Ackerman, 1999). On this model, individuals engage in corruption and crime because they expect the benefits to outweigh the risk of being caught and the penalty if they are punished. The policy solution that follows from this model focusses on changing the incentives facing actors so that they no longer find it worthwhile to engage in corruption. Law enforcement can be a key part of this, since effective law enforcement increases the risks of being caught, while tougher regimes can also increase the sanctions associated with prosecution (Becker, 1968). Indeed, regulation that seeks to prevent

money laundering and provides the police with powers to recover wealth and assets accrued through corrupt and criminal activity is explicitly designed to 'take the profit out of crime' on the assumption that this will make corruption and criminal enterprise less attractive.

An alternative model of human behaviour is grounded more in sociology, and suggests that humans tend to act in ways that follow the rules or norms within so-called 'logics of appropriateness' (March & Olsen, 2004). On this model, individuals do not make calculations before acting, they simply do what everyone else in the same position does – they follow a norm, and if the norm within a certain profession or community happens to involve corruption or criminality, they engage in such acts not on the basis of any calculation but just because everybody else does. On this model, fighting corruption and criminality requires changing the norms that prevail in certain groups or that pertain to particular roles – or, potentially, persuading people that norms associated with one of their roles are more important or salient than rules associated with another role. For example, for public officials and police officers who face opportunities to engage in corruption, this might mean convincing them to prioritise impartiality in public office over a community norm to (ab)use your access to power to help friends or re-pay allies.

The sociological model resonates with recent critiques of the economic approach which note that, in contexts where corruption is systemic, the incentives approach is unlikely to be successful owing to a collective action problem. If you expect that others will behave corruptly, including the law enforcement officers and regulators who are supposed to detect and punish your behaviour, then the perceived risks and costs of being caught plummet. Even if an individual in this context wishes to fight corruption, they might well expect that others will not, meaning that they will be the only one incurring the social costs of eschewing corrupt opportunities and the efforts of one individual will achieve nothing in terms of changing the situation. If all individuals calculate in the same way, then anti-corruption efforts in situations of systemic corruption will fail (Persson, Rothstein, & Teorell, 2013).

Marquette and Peiffer point out that the two approaches are not alternatives but rather both have explanatory value, and equally, anti-corruption strategies are more likely to be successful if they build on both models (Marquette & Peiffer, 2015). Our theory of change for evaluating CACP therefore develops two theoretically-informed pathways to impact, one which looks at how the interventions change incentives and disincentives and another which analyses how interventions resonate with or challenge local norms that support corrupt behaviour.

ACAs are publicly-funded entities entrusted with the objective of fighting corruption and reducing the opportunities for corruption by means of preventive and/or repressive measures (de Sousa, 2010). They are a typical example of an anti-corruption intervention that is rooted in an economic approach to understanding human behaviour. Most ACAs have a law enforcement role: they are involved in investigating corrupt acts that involve violations of the law, and some ACAs also have prosecutorial powers (although others pass on cases to prosecutors) (Heilbrunn, 2004). As such, their aims are to detect and punish crime, thereby – on the economic logic - increasing the risks and penalties and therefore deterring corruption. However, ACAs do not work in a vacuum, and the political and societal context in which they operate may well be one of systemic corruption, meaning that they cannot rely on the integrity of *all* members of their own institution, nor members of other organisations in the accountability ecosystem. Their success may also depend on changing logics of appropriateness that foster corruption and criminality; this requires attention to impact pathways based on the sociological model, requiring an understanding of how the work of ACAs might challenge old norms and start to build new ones.

Each evaluation tool developed here considers both the economic and sociological pathways to change and the interactions among them. Thus, our assessment of organisational leadership highlights the importance of champions of reform for changing norms as well as incentives. Our network analysis tool builds on research about the value of social networks for coordination among

law enforcement professionals (LaForge, 2017), but argues that the importance of networks lies not only in increasing the risk of detection, but also in creating a community of professionals who gain strength from shared commitment to change in the face of challenging circumstances and are thus more empowered to overcome the collective action problem. The learning shared here is applicable to other programmes that tackle corruption by supporting law enforcement institutions and highlights the benefits of a partnership-based approach.

Measurement challenges

Scholars have struggled with how to measure corruption for the last three decades. The first generation of measures sought to assess the level of corruption in an entire country and was typically based on surveys of public and expert perceptions; Transparency International's Corruption Perceptions Index, which has since 1995 published annual country scores and rankings, is the prime example. These measures were criticised primarily for being aggregate measures, potentially masking important variations in types of corruption in different areas or sectors; and for being based on subjective perceptions, which are prone to various kinds of bias (Andersson & Heywood, 2009; Heywood & Rose, 2014; Philp, 2006).

A second generation of measures sought to address these problems (Johnston, 2000; Reinikka & Svensson, 2006; Sequeira, 2012): for example, Transparency International launched an index of 'bribe payers' in business while the World Bank's Business Environment and Enterprise Performance Survey (BEEPs) measured businesspersons' experience of corruption rather than their perceptions. Nevertheless, the measures still relied on individuals to report accurately; this may lead to under-reporting, either because of fears associated with reporting behaviour that is considered 'deviant' or, conversely, because certain informal behaviours are seen as so normal that they are not considered corrupt.

A third generation of corruption measurement relies not on self-reporting but on proxy indicators (Johnson & Mason, 2013). For example, by collecting detailed data about public finances, and identifying gaps between revenue and spending, leakage can be identified that may at least partly reflect corruption. Another method analyses 'big data' on public procurement processes to construct an index from 'red flags' – irregularities that may indicate that the procedure has been corruptly manipulated to steer contracts towards favoured bidders (Fazekas, Tóth, & King, 2016). The advantage of such proxy indicators is that they use objective data. One disadvantage is that the irregularities they identify may have numerous causes, only some of which relate to corruption. As such, proxies are best used to spot patterns in large datasets, which can then be investigated to determine whether corruption is present, rather than to *measure* corruption in a particular context.

Ideally, an evaluation of the impact of an anti-corruption intervention would measure corruption levels before and after – as well as with and without - the intervention, as well as determining the contribution of the intervention to the observed change. This is always challenging for evaluators, given the complexity of social environments in which policies and programmes are implemented. In the case of corruption, even the first task – measuring corruption levels before and after – is difficult, and the task of attributing causality is likely to be especially complicated.

Moreover, most ACAs have wide remits and seek to achieve macro-level or systemic change (de Speville, 2010; Choi, 2011); this is certainly true of the agencies supported by the CACP. In theory then, success might be reflected in aggregate first-generation corruption measures, at least in the medium term. Yet perceptions of corruption, which are the focus of many survey-based measures, might also be influenced by ACA actions in counterproductive ways. Imagine a context where corruption has been allowed to occur almost unchecked under a certain political leader, but an incoming leader invests resources in an intensive effort to crack down on corruption. One effect of the crackdown is that the public becomes far more sensitised to corruption than it used to be. In a perceptions-based survey, this might show up as an increase in perceived levels of corruption, even

if the crackdown is in fact reducing the incidence of corruption. Perceptions are influenced by key events and circumstances, and higher-profile ACA action might well influence perceptions of corruption in ways that show up unhelpfully in the statistics. Similarly, successful ACA work might lead to more prosecutions for corruption; this would translate into an increase in recorded corruption crimes, which might be misinterpreted as an increase in corruption levels, when in fact it would represent a higher share of corrupt activity being prosecuted.

Evaluations of ACAs must go beyond macro-level indicators of corruption and need to take account of the context, what is achievable in that context, and how the work of the ACAs contributes to changes in attitudes and behaviour. Isolating the contribution of a particular ACA or a certain donor programme is also difficult given the complex range of factors which contribute to patterns of corruption in a country - including the socio-political context; institutional arrangements; legislative framework and accountability mechanisms.

In the case of the Caribbean, this context included a complex relationship between corruption and organised crime, the presence of multiple donors with overlapping objectives and, in the case of the RSS-ARU, a mandate that covers several jurisdictions, each with their own political and social dynamics. Although Jamaica is an upper-middle-income country, the unemployment rate is around 12% and young men in particular are vulnerable to being recruited into gangs. Law enforcement agencies lack capacity to investigate and prosecute crime, but are sometimes perceived to be part of the problem. This lack of trust hinders the investigative work of the police – for example, because individuals are unwilling to serve as witnesses and may see local gangs as offering more protection than the police. This creates an unhelpful cycle, whereby the task of law enforcement agencies becomes progressively more difficult, they are perceived to be less effective, and confidence in the rule of law is further eroded.

In such complex environments, focusing on capacity as a core part of the evaluation of ACAs can be a way of defining measurable intermediate outcomes, whilst situating these within a theoretical framework which ensures that they support the achievement of longer-term goals. In the Caribbean, for example, it was very clear that the effectiveness of the organisations evaluated was a function not only of their own internal capacity but also of their ability to harness the capacity and will of many other administrative and judicial bodies on which they rely to turn investigations into successful prosecutions. Network connections among a wide range of institutions help to build a perception that the enforcement net is tightening, as well as allowing for the diffusion of new anti-corruption norms, to replace old unhelpful norms of tolerating corruption or wishing to avoid being seen as a snitch for helping the police.

Theory-based evaluation

Theory-based evaluation is not a single, clearly defined approach. Many terms are used to describe it, often interchangeably, and without consistent definitions (Funnell & Rogers, 2011; Rogers, 2007).[6] Consequently, there is no single shared understanding of what a good theory-based evaluation looks like, or how evaluators should go about conducting one (Coryn, Noakes, Westine, & Schroter, 2011). For the purpose of evaluating capacity on anti-corruption programmes, our understanding of a theory-based approach is one in which the evaluation design and application is explicitly guided by theory about how a programme leads to change (Chen, 2012; Coryn et al., 2011; Morris & Fitz-Gibbon, 1996); in this case, theory and research on what works in fighting corruption are relevant.

While ACAs are often seen as an essential element of a national anti-corruption framework and their use has proliferated considerably in the last two decades, research suggests that their success depends on a range of factors, including the extent of their mandate, the degree of independence and

[6] e.g. theory-based, theory-driven, theory of change, intervention theory, outcomes hierarchies, programme theory, programme logic – see Funnell & Rogers, 2011; Rogers, 2007)

resources they are given, and the extent of staff professionalism and skill, as well as rule of law in the given context (Bolongaita, 2010; de Sousa, 2010; Doig et al, 2006; Heilbrunn, 2004). Some of these factors are beyond their control. It is therefore necessary to understand the context in which ACAs operate and to identify which aspects of that context are within their scope of control, influence and concern. A political economy analysis is critical, for example, to understanding whether the political space exists for them to fulfil their mandate, and how other actors are likely to respond to ACA operations. ACAs are also inherently constrained by factors such as resources (human and financial) and the nature of their mandate, which varies considerably. Understanding causality thus requires careful observation of changes as well as understanding mechanisms so as to determine the contribution of the intervention. Attribution was challenging in the cases studied here because there was no possibility to build or explore a counterfactual scenario in which no intervention was implemented; additionally, the intervention was one of a myriad of factors influencing change in the period examined, making it difficult to isolate direct and indirect effects.

However, a theory-based approach enabled the evaluation to unpack how, why and to what extent ACAs in the Caribbean do or do not achieve outcomes and which contextual factors influence their effectiveness. The evaluations focused on assessing whether the interventions were one of the causes of observed changes. To ensure that we applied the theory in an appropriate and relevant way, the evaluations were conducted in a highly participatory manner, engaging all of the key stakeholders in theory of change workshops to ensure that theoretically-informed pathways to impact resonated with the evaluands. Working together to elaborate the mechanisms and linkages between interventions and expected changes fostered consensus around how the programme was expected to work and what needed to be evaluated. The reconstructed theory became the scaffolding for the evaluation which ensured that all components and assumptions of the programme were considered. A Contribution Analysis approach (Mayne, 2008) was applied, elaborating the theory in stakeholder workshops, gathering empirical evidence - through document review, Key Informant Interviews (KIIs), Focus Group Discussions (FGDs) - and testing the causal claims/linkages through triangulation and use of analytical tools, and finally refining the theory and evidence narrative about what had worked (or not) and why. The evaluation deployed three analytical tools - capacity assessment, network analysis and policy tracker - to understand causality and to provide evidence about the interventions' contribution to observed changes. Each is discussed in detail in the following sections.

Tools for Evaluating Capacity in ACAs

Capacity Assessment

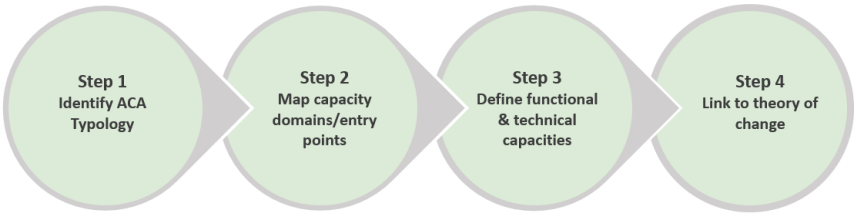
Evaluating capacity through external institutional assessments can help identify areas of strength and weakness, although it should not replace consistent monitoring and reflection by institutions themselves (Schütte, 2017). Capturing capacity at a 'point in time' can provide important lessons on what works and what does not, to inform course correction mid-programme cycle.

A Four-step Methodology

The evaluations designed a four-step methodology (see Figure 1), in collaboration with the ACAs, which envisaged capacity as distinct but interconnected domains. This methodology is an adaptation of the UNDP's Capacity Assessment for Anti-Corruption Agencies (United Nations Development Programme, 2011) and the UK DFID's Building Capacity to Use Research Evidence programme, which identifies interconnected 'levels of change' for capacity development (Punton et al., 2016). A benchmarking tool was designed, enabling a structured assessment within functional and technical areas, with a four-point rating scale. Context is important particularly since progress towards change is not necessarily linear and there is a need for adaptive capacity in an evolving context.

Step 1. ACA Typology

Figure 1. ACA Capacity Assessment: A Four-Step Methodology



UNCAC Articles 6 and 36 set basic principles for ACAs,[7] but adoption of the articles at the country level varies, with functions and mandate influenced by the political, administrative and resource environment. In order to build an adaptive capacity framework that considers effectiveness, it is necessary to categorise ACAs according to their constitutional mandate, level of independence, oversight and powers. In the case of Jamaica, there are three ACAs with varied functions and mandates dispersed across the administrative system; in the Eastern Caribbean, there is one regional entity focussed on asset recovery and financial investigation that has a preventative function across the OEC member states (see Table 1). The primary intervention focus of CACP is to build their capacity.

Table 1. The supported institutions and their mandates

Institution	Mandate	Enabling law/decree	Jurisdiction
Major Organised Crime Agency (MOCA)	Investigates and disrupts corruption and organised crime perpetrated and facilitated by Politically Exposed Persons (PEPs), police officers and High-Value Targets (HVTs).	Legislation (the MOCA Bill) to establish MOCA as an entirely independent entity passed the Senate in June 2018; it will be fully operationally independent once the House of Representatives passes the final regulations.	Jamaica
Independent Commission of Investigations (INDECOM)	Responsible for undertaking investigations concerning actions by members of the security forces and other agents of the state that result in death or injury to persons or the abuse of the rights of persons.	Independent Commission of Investigations (INDECOM) Act (2010).	Jamaica
Financial Investigations Division (FID)	Designated authority to address issues pertaining to financial crime, specialised in	The Proceeds of Crime Act 2007 (2007); the Financial Investigations Division Act	Jamaica

[7] Similar institutions such as Financial Intelligence Units apply the Egmont Group Statement of Purpose and Principles of Information Exchange; Audit Agencies (INTOSAI Standards for external government auditing); National Human Rights Institutes (Paris Principles by the ICC for NHRIs).

	anti-money laundering and asset recovery, and prime contact for Financial Action Task Force.	(2010) and the Terrorism Prevention Act (2005). It is situated within the Ministry of Finance and Public Service.	
Regional Security System – Asset Recovery Unit (RSS-ARU)	Supports financial investigations and asset recovery activity in jurisdictions.	The RSS-ARU is governed by the regional Security System Treaty (1996).	Barbados, St Lucia, St Vincent & Grenadines, St Kitts & Nevis, Antigua & Barbuda, Grenada, Dominica.

Step 2. Four distinct but interconnected capacity domains

The OECD-DAC defines capacity as "the ability of people, organisations and society as a whole to manage their affairs successfully"(OECD, 2010; p1.). Increasingly, capacity is understood holistically, as multiple layers of relationships and processes, incorporating also the concept of resilience or sustainability—the ability of a society or sector to continue to develop important skills, behaviours, networks and institutions into the future (Kaplan, 1999; Ubels and van de Gronden., 2010, quoted in Punton et al., 2016). For the evaluation, capacity 'entry points' or domains were identified through discussions with the ACAs, leading us to classify capacity in four distinct but interconnected domains: Institutional; Interpersonal; Organisational and Individual (Figure 2). This enabled the evaluation to unpack where components interact in dynamic ways, allowing for systemic rather than single component thinking better able to accommodate complex and non-linear trajectories.

Figure 2. Four Domains of Capacity

Step 3. Defining functional and technical capacities: building the Benchmark Tool

Within these four domains, ACAs have two particular types of capacities – functional and technical (UNDP, 2011). Functional capacities are necessary for managing ACAs and are linked to the core

issues and challenges which in some cases were surfaced through step 1. Technical capacities are associated with particular areas of professional expertise which are linked to the specific function and mandate of ACAs. UNDP suggests that while conventional capacity development interventions have focused largely on technical capacities, there is increasing recognition that enhancing functional capacities offers stronger sustainability and contribution to outcomes.

Building on UNDP guidance, a benchmarking tool was co-created with the ACAs to enable dimensions of capacity to be assessed in a standardized way. This enables the programme to use the findings as a 'point in time' analysis while the ACAs themselves can incorporate the tool into their work planning to ensure that capacity is strengthened on an ongoing basis. We identified fifteen functional and technical capacities (with seven emphasised as key areas of focus[8] for this evaluation and for the programme-level synthesis), allowing us to capture the diversity of these four ACAs. For each type of capacity, benchmarks corresponding to a Likert 1-4 scale (none-basic-moderate-high) were refined to correspond to the Jamaican and Eastern Caribbean context, recognizing that capacity development is an inherently political and complex process where outcomes do not necessarily evolve in a controlled and linear fashion (UNDP, 2007).

Step 4. Link to Theory of Change

Across the four ACAs, the evaluation conducted 39 key informant interviews, six focus group discussions (structured by management level) and an online staff survey that achieved a 55 per cent response rate. Data were gathered in a structured way, ensuring that each data collection tool had a clear link to the respective capacity domain and was analysed against the benchmarking criteria. Each ACA was then rated by the evaluation team against the four capacity domains (with scores verified by the ACAs), applying a four-tier traffic light system[9] for each criterion. Synthesis ratings were then derived from the individual assessments, which enabled the evaluation to assess the contribution of the intervention to the observed changes at the respective levels of control, influence and concern on the reconstructed theory of change. Ratings were synthesised across the corresponding seven capacities identified as key areas of focus for this evaluation (see step 3), and an average rating was applied (across the agencies) which formed the programme-level rating.

The framework assessed capacity in a standardised but holistic way, establishing a midpoint baseline for the intervention that can be repeated at the end of the programme cycle. This approach has some limitations: the ratings were verified at one point in time and were subject to individual interpretation of the data, albeit based on careful coding and triangulation by a team of evaluators. On the other hand, the framework is a bespoke measurement tool for these particular agencies, the co-creation of which enhanced learning and reflection. The individual agency assessments also provided a basis for the creation of agency action plans.

In terms of findings, the first evaluation found that in the Jamaican context, overall anti-corruption capacity was constrained by the dispersion of responsibilities and mandates for anti-corruption functions across three institutions (MOCA, FID and INDECOM), particularly since two of them – MOCA and FID – were yet to gain functional and budgetary autonomy (although steps were underway to improve this situation). However, this structural weakness was mitigated by strong inter-agency relationships which helped to ensure that things got done. The evaluation noted that these relationships were largely informal and recommended greater formalisation of inter-agency relationships to make them more stable and less vulnerable to being broken by turnover in personnel.

At the individual level, the core and functional capacities identified by the ACAs were scored as moderate, with a need for continued upskilling and training in new techniques. The evaluation found little evidence of a strategic approach to learning and development and hence this formed the basis

[8] Seven key areas of focus: Institutional (1. ACA goals are set internally and aligned to national plans); Interpersonal (2. Inter-agency coordination/Network); Organisational (3. Dependence of Head of Agency; 4. Human resource planning and performance; 5. Operational policies and processes; 6. Promotion of integrity – public relations); Individual (7. Capacity to deliver core functions).

[9] The four-tier traffic light system uses an adapted version of the Independent Commission for Aid Impact ratings system available: Here.

for another recommendation. However, most of the individual leaders of the organisations were rated very highly, both by their staff and by peers, creating a strong 'tone at the top' which helped to inspire more junior members of the organisation. These agency leaders seemed to play a critical role in motivating staff, who often faced great personal risk and even stigma, particularly in 'small island' contexts where officers and perpetrators often have social network ties. Inspirational leadership was critical in building new norms of professional integrity and pride, while training was also valued as a way of signalling professionalism and expertise. Across the ACAs, the weakest area was organisational capacity, which reflected some common issues: for example, dependency on leadership; resource constraints; weaknesses in Human Resources planning; inefficiencies in processes (including strategic alignment and internal monitoring); and a lack of capacity to strategically engage and communicate.

Network analysis

Theory suggests that the effectiveness of ACAs may be heavily dependent on their ability to coordinate. According to the economic theory of anti-corruption, coordination improves the capacity to detect crime and collect evidence; on the sociological theory, the creation of network ties among anti-corruption professionals helps overcome the expectation that others will not act and builds new professional norms of integrity and courage in the face of adversity. This is supported by empirical research on preventive measures, criminalization, international cooperation and asset recovery which demonstrates that tackling corruption requires the support and engagement of many institutions (laForge, 2017). In the case of Jamaica, interpersonal links across the ACAs are arguably as important as the formal institutional arrangements among organisations.

Network analysis focuses on the relationships between actors within a system, represented in a visual map of the individuals/organisations (nodes) and their relationships (links, ties or edges). For the ACAs in Jamaica, the evaluation built on an initial stakeholder mapping by documenting details about the density and quality of relationships among actors. We used self-reporting in KIIs and FGDs to identify the nodes (actors) and the ties of cooperation among them, as well as attributes of the nodes and relevant qualities of the ties (e.g. formal vs. informal). The nodes and links in the network were documented and mapped using visualisation software, Polinode. The map (see figure 3) illustrates where actors are strongly connected, where ties are missing and allows us to identify 'gate-keepers' which link different parts of the network. More qualitative dimensions, such as the motivations for the relationships were elaborated in narrative form in the full evaluation.

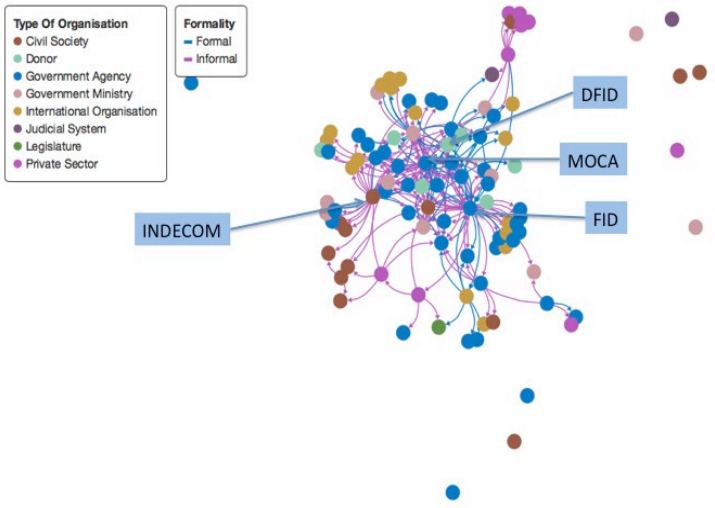


Figure 3. ACA Network Analysis

The network analysis resonated with the ACAs, which had identified inter-agency coordination and the building of multi-disciplinary teams as key short-term outcomes of the CACP and a crucial dimension of capacity. It revealed that the agencies had more inward – within CACP - than outward connections, highlighting a need to engage in more externally-facing communications. On the economic logic, demonstrating and promoting their activity to external audiences would help to persuade the public that the risks of crime and corruption being detected are increasing, helping to deter deviant behaviour and make cooperation with the authorities seem more worthwhile. On the sociological logic, greater communication would also help to create a sense that the norms are changing and that collective action against corruption is becoming more feasible.

However, given the varied mandates of the different institutions, there are also differences in the type and density of external networks that are optimal for operational effectiveness. In a context in which there is a widely dispersed mandate to tackle organised crime and corruption, there are risks that resources will be stretched and this will reduce effectiveness overall. Moreover, there are many possible entry points for political interference. The ACAs' ability to function and, in particular, to tackle 'Big Fish', is heavily reliant on their ability to build close, trusting and collaborative networks and to coordinate action among a wide range of organisations. Our network analysis showed good informal relationships among the key ACAs, but also highlighted that civil society organisations remain largely on the periphery: intervention is more top-down than bottom-up. Some government ministries are also fairly peripheral. This may reflect the political sensitivity of fighting high-level corruption, which means that ACAs need to be cautious in sharing details about their work, but future progress is likely to be partly dependent on building more allies in these circles.

Policy tracker

The ACAs supported by CACP are sometimes hindered in their work by weaknesses in the governance framework or bottlenecks elsewhere in the accountability system, e.g., backlogs in the courts, or the absence of secondary regulation needed to implement laws on the statute books. In other cases, their investigations reveal loopholes in regulatory frameworks that create opportunities for corruption and fraud. If ACAs identify such governance gaps, they can also advocate for reform aimed at closing off opportunities for crime, ensuring that an activity is better resourced or a procedure elaborated, or increasing their own powers in a particular area. This area emerged as an important activity for all of the ACAs which we evaluated, and one which is critical to their long-term success, yet it was often not explicitly set out in work plans and tended rather to be dealt with on an ad hoc basis. This prompted the evaluation team to develop a tool for measuring their progress in advancing policy advocacy goals.

The policy tracker was adapted from a tool for assessing policy advocacy work originally developed by USAID and utilised to evaluate the ACAs' efforts to influence the legislative and regulatory environment in which they operate (Hirschmann, 2002). We enhanced the tool to focus on the relationship between change in policy and change in norms, in line with theory about norm diffusion (Finnemore and Sikkink, 1998). We identified six stages of policy change from concept to full implementation, against which progress could be tracked (see Figure 4 below).

Figure 4. Policy scale (tracker)

The first step is a change in discourse. This occurs when a few policy entrepreneurs manage to get the issue on the agenda of a relevant stakeholder group – whether this is the general public or, more often, a relevant ministry or parliamentary committee. This may lead to a new policy proposal being developed (stage 2) and put before parliament or, in the case of secondary legislation or regulatory change, actioned by the relevant part of the bureaucracy. The proposal may then be accepted - or amended - and is finally 'adopted' (stage 3), if it does not fail or simply stall. Even once adopted, it is not certain that it will be 'implemented' (stage 4), since this often requires engagement from a variety of actors who may not be aware of their new powers or may be reluctant to use them. 'Enforcement' (stage 5) signifies not just that a new law or regulation is implemented, but also that compliance is monitored and non-compliance punished. Stage 6 is reached when the new behaviour becomes habitualised, i.e., there is a 'change in culture'.

While the policy tracker was helpful in identifying advocacy goals and monitoring progress, evaluating partners' achievements in this regard was difficult. First, the policy initiatives differed considerably in scope; second, the 'starting points' varied somewhat depending on the issue; third, it is difficult to assess the partners' contribution to the progress achieved, given that public policy is made in a complex environment with many interest groups and influences. As such, it is not appropriate to draw conclusions about *effectiveness* in this area. Rather, the policy tracker allowed us to record the ACAs' achievements in a previously neglected part of their work and established a benchmark which they can use to plan, track and evaluate progress in future.

Conclusions and lessons

On evaluating ACAs

These evaluation tools recognise that ACAs face a challenging and complex task and tend to operate in environments of weak governance and low resources. A theory-based approach enhances the utility of an evaluation in such contexts by providing a basis for key stakeholders to re-design interventions and re-allocate resources when political opportunities arise or, alternatively, channels of influence are closed off. A participatory approach to the theory of change is also critical in such a highly sensitive area. Most individuals working in these agencies face extreme risk and social stigmatisation, meaning that trust and respect between evaluator and evaluand is of the utmost importance.

The tasks of reducing corruption and organised crime so as to achieve better governance are long-term goals, likely to require significant investment in institution-building and the nurturing of change in norms and behaviour over generations. Moreover, while technical advances in tracing financial transactions increasingly allow investigators to track networks of criminals, including politicians in positions of power, catching and prosecuting 'big fish' requires considerable agency capacity and coordination over long periods of time. It is also a high-risk activity that may prompt retaliation by powerful targets unless there is sufficient political will and commitment in government, at least to allow agencies to operate independently.

Given these constraints, we suggest that programme evaluations of ACAs are fairer and more constructive if they focus on medium-term goals, where capacity building is key. To understand what kinds of capacity are important, we worked with stakeholders to elaborate a theory of change

that is grounded in the two main strands of anti-corruption theory, and which considered relevant features of the social and political context. Because research suggests there is a need to change norms as well as incentives, we developed a strong section on organisational leadership in our organisational capacity assessment tool. Because research suggests that social networks among law enforcement professionals are critical to initiating the kind of collective action necessary to conduct anti-corruption investigations and bring cases, we developed a tool to analyse formal and informal networks among key partners. Because long-term change requires a dynamic and adaptive approach to fighting corruption (because corrupt and criminal practices evolve rapidly), we created a policy tracker to collect evidence relevant to the sustainability of anti-corruption efforts in challenging contexts. Our focus on these tools recognises that the success of ACAs will depend on a number of contextual factors that are beyond their control, but when political opportunities arise, we argue, ACAs will be best able to utilise them if they have built capacity on the dimensions discussed here.

On how to design ACAs

Our experience in these evaluations and other work has also yielded lessons relevant to the work of ACAs elsewhere.

First, despite a global trend of investing in law enforcement as a key instrument for addressing corruption, such approaches tend in practice to focus more on organised crime and on post-corruption offences such as money laundering than on grand corruption itself. While this is not surprising, because there is less experience of investigating grand corruption and such investigations require a willingness to confront those in power head on, in a situation where organised crime is highly interrelated with grand corruption, such an approach risks undermining the overall goal by creating a sense of impunity and may mean that short-term achievements prove unsustainable. More broadly, there is a need to ensure that law enforcement work is complemented by work to address root causes, including youth unemployment and inequality, as well as strategic planning for unexpected, negative consequences of law enforcement activity.

Second, law enforcement approaches should be accompanied with efforts to displace old norms and diffuse new anti-corruption norms. While learning in international development increasingly recognises the need to pay attention to the political and social context, programming related to ACAs often reverts to a more traditional law enforcement strategy which is fundamentally based on the economic accounts of human behaviour and seeks to change incentives. This work in the Caribbean reveals how contextual factors - including the sporadic and inconsistent nature of the Jamaican government's attention to fighting major corruption and organised crime, and the perceived prevalence of deeply embedded social norms that facilitate corrupt practices such as a 'small island' culture - are inherent to defining the problem and designing tailored solutions. This suggests a need to accompany law enforcement approaches with efforts to diffuse new anti-corruption norms. This should be regarded as an integral part of the work of ACAs generally, and not seen as an activity separate to law enforcement work.

Third, building networks of expertise is critical to addressing complex cases, since it allows for information-sharing across administrative mandates and across geographical borders, mimicking the nature of much criminal and corrupt activity. Such networks are often facilitated by informal or individual relationships, and this may particularly be the case in 'small island' contexts such as the Caribbean. External personnel providing technical assistance need to be able to integrate with and earn the trust of local networks, which again often relies on building informal relationships. Many of the CACP experts providing technical assistance have worked in Jamaica for several years, building their reputation and enjoying an exceptional degree of trust-based access to key stakeholders. Yet this may not be straightforward to replicate in other contexts. Moreover, to build robust institutional capacity, it is necessary to formalise relationships among institutions with memoranda of understanding and administrative protocols for cooperation. Even though informal

relationships are working well in the Caribbean context, relying on these alone makes structures vulnerable to future political interference and changes in leadership.

References

Andersson, S., & Heywood, P. M. (2009). The Politics of Perception: Use and Abuse of Transparency International's Approach to Measuring Corruption. *Political Studies*, *57*, 746–767.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. In *Journal of Political Economy* (Vol. 76). https://doi.org/10.1086/259394

Bolongaita, E.P. (2010) 'An exception to the rule? Why Indonesia's anti-corruption commission succeeds where others don't: a comparison with the Philippines' ombudsman', U4 Anti-Corruption Resource Center, August 2010.

Chen, H. T. (2012). Theory-driven evaluation: Conceptual framework, application and advancement. In R. Strobl, O. Lobermeier, & W. Heitmeyer (Eds.), *Evaluation von Programmen und Projekten für Eine Demokratische Kultur* (p. 226). https://doi.org/10.1007/978-3-531-19009-9

Choi, J. (2011) Measuring the Performance of an Anticorruption Agency: The Case of the KPK in Indonesia, International Review of Public Administration, 16:3, 45-63, DOI: 10.1080/12294659.2011.10805207

Coryn, C. L. S., Noakes, L. a., Westine, C. D., & Schroter, D. C. (2011). A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. *American Journal of Evaluation*, *32*(2), 199–226. https://doi.org/10.1177/1098214010389321

Doig, A., D. Watt, and R. Williams. (2006). Hands-on or hands-off? Anti-corruption agencies in action, donor expectations, and a good enough reality. *Public Administration and Development*, 26/2: 163-172.

De Sousa, L. (2010) Anti-corruption agencies: between empowerment and irrelevance. *Crime, Law and Social Change*, 53/1: 5-22.

De Speville, B. (2010) Anticorruption Commissions: The "Hong Kong Model" Revisited, Asia-Pacific Review, 17:1, 47-71, DOI: 10.1080/13439006.2010.482757

Fazekas, M., Tóth, I. J., & King, L. P. (2016). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal of Criminal Policy and Research*, *22*(3), 369–397.

Finnemore, M., & Sikkink, K. (1998). International norm dynamics and political change. *International organization*, *52*(4), 887-917.

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco: Jossey-Bass.

Heywood, P. M., & Rose, J. (2014). "Close but no Cigar": the measurement of corruption. *Journal of Public Policy*, *34*(03), 507–529. https://doi.org/10.1017/S0143814X14000099

Heilbrunn, J.R. (2004). Anti-corruption Commissions: Panacea or real medicine to fight corruption? World Bank Institute, Washington DC.

Hirschmann, D. (2002). "Implementing an indicator": Operationalising USAID's "Advocacy Index" in Zimbabwe. *Development in Practice*, *12*(1), 20–32. https://doi.org/10.1080/09614520220104257

Johnson, J., & Mason, P. (2013). The Proxy Challenge: Why bespoke proxy indicators can help solve the anti-corruption measurement problem. In *U4 Brief*.

Johnston, M. (2000). *The New Corruption Rankings: Implications for Analysis and Reform*.

Kaplan, A. (1999). The development of capacity. Geneva: Non-Governmental Liaison Service. No.

658.404.K17d)

LaForge, G. (2017). *The sum of its parts: Coordinating Brazil's Fight Against Corruption, 2003 – 2016*.

March, J.G., and J. P. Olsen. (2004) 'The logic of appropriateness'. *The Oxford Handbook of Political Science.* Oxford: Oxford University Press.

Marquette, H., & Peiffer, C. (2015). Corruption and Collective Action. *Development Leadership Program, U4 Anti-Corruption Resource Centre*, 22.

Mayne, J. (2008). Contribution analysis: An approach to exploring cause and effect. *ILAC Brief*.

Morris, L. L., & Fitz-Gibbon, C. T. (1996). Theory-Based Evaluation. *Evaluation Practice*, *17*(2), 177–184.

OECD. (2010). *Capacity Development : a Dac Priority*. (November), 2010–2011.

Persson, A., Rothstein, B., & Teorell, J. (2013). Why Anticorruption Reforms Fail—Systemic Corruption as a Collective Action Problem. *Governance*, *26*(3), 449–471. https://doi.org/10.1111/j.1468-0491.2012.01604.x

Philp, M. (2006). Corruption, definition and measurement. In *Sampford, C., Shacklock, A., Connors, C., and Galtung, F (eds.) Measuring Corruption (2006), pp. 45–56. Aldershot, UK: Ashgate* (pp. 45–56).

Punton, M., Brown, C., Echt, L., Fisher, C., Lloyd, R., Morton, S., & Vogel, I. (2016). How Can Capacity Development Promote Evidence-Informed Policy Making? *Literature Review for the BCURE Programme*, (April).

Ravallion, M. (2009). Evaluating three stylised interventions. *Journal of Development Effectiveness,* 1/3: 227-236.

Reinikka, R., & Svensson, J. (2006). Using Micro-Surveys to Measure and Explain Corruption. *World Development*, *34*(2), 359–370. https://doi.org/10.1016/j.worlddev.2005.03.009

Rogers, P. J. (2007). Theory-Based Evaluation: Reflections Ten Years On. *New Directions for Evaluation*, *114*, 63–81. https://doi.org/10.1002/ev

Rose-Ackerman, S. (1999). *Corruption and Government: Causes, consequences and reform*. Cambridge University Press.

Schütte, S. (2017) *Bespoke monitoring and evaluation of anti-corruption agencies*, U4 Brief.

Sequeira, S. (2012). Advances in Measuring Corruption in the Field. In D. Serra & L. Wantchekon (Eds.), *New Advances in Experimental Research on Corruption* (pp. 145–176). Bingley, UK: Emerald Group Publishing Limited.

Ubels, J. and J. van de Gronden. (2004)From theory to practice: case studies in meso-level capacity-building. *Capacity Organisation* 22/2-3

United Nations Development Programme. (2011). *Practitioners Guide: Capacity Assessment of Anti-Corruption Agencies*.